Scalable, High-performance, Elastic Software OpenFlow Switch in Userspace for Wide-area Network

Yoshihiro Nakajima, Tomoya Hibi, Hirokazu Takahashi, Hitoshi Masutani,
Katsuhiro Shimano, Masaki Fukui
Network Innovation Laboratories, Nippon Telegraph and Telephone Corporation
e-mail: nakajima.yoshihiro@lab.ntt.co.jp

Abstract

This paper presents the design and implementation of scalable, high-performance, and flexible software-based OpenFlow 1.3 switch in userspace on commodity Intel x86 PC servers for wide-area network. To achieve the high-performance and high-scalability in flow processing, the software-based switch leverages the state of the art of multi-core CPUs and OS technology for high-speed network I/O. The switch achieves 10.1MPPS with 100K flow rules and short packet (64B) in case of a network gateway between a VLAN-based data center network and IP-VPN/MPLS with Intel xeon E5-2660.

1 Introduction

Recent advances of server-virtualization technology and cloud computing technology allows many services, such as web, database, and application, run with virtualized computing resources in a data center. A large-scale data center runs more than thousand tenant network with tens of thousands virtual machine, thus, over hundreds of thousands flows must be handle in network including both hardware switch and software switch. On the other hand software-defined networking (SDN), OpenFlow, and network functions virtualisation (NFV) attract network carriers and service providers because SDN technology may realize service flexibility and rapid adaptation for service differentiation and cost reduction. NFV technology may accelerate service node migration from the special purpose network function nodes to software-based packet processing with commodity servers for flexible management, deployment and delivery. Software switch becomes considerably important from the scalability and performance for cloud computing, SDN, and NFV. Existing software OpenFlow switches [2], such as Open vSwitch [3], LINC [], focus on data center networking and virtual machine, therefore, only limited protocols are supported. In addition, packet forwarding/processing and flow rule scalability are also limited from the NFV point of view. High-performance software-based packet processing techniques have been studied for over decades [5, 1], however, they are limited to optimization and acceleration in a specific layer, not in inter-layer. This research aims to provide high-performance software OpenFlow switch to realize 10-Gbps-wirerate with 1M flow rules using commodity servers for various network domain.

2 Design

We summarize the requirements of Software-based OpenFlow switch for wide-area network by reference to report of SDN experience in a network carrier [4] as follows:

- A switch must run on the commodity pc server and commodity network interface card (NIC).
- A switch should provide a gateway function to allow connect different various network domains, such as DC, IP-VPN, MPLS and access NW.
- A switch can achieve 10Gbps-wire rate with more than 1M flow rules with low-latency packet processing and flexible flow lookup using multiple-tables.
- A switch can run in userspace and decrease the tight-dependency to OS kernel code for easy software upgrade and deployment.
- A switch should support various management and configuration protocols.

Figure 1 shows the design overview of the software switch. The switch consists of two main components, agent and data-plane. The agent provides a unified data store functionality of switch resource configuration and management. These protocols and interfaces, such as OpenFlow-wire, OF-CONFIG, OVSDB, CLI, SNMP, are implemented as a module to the data store of agent. This framework allows a developer to implement easily new configuration and management protocol and their extensions. The data-plane exploits userspace network I/O library to run the switch in userspace. Control messages and event messages between agent and data-plane

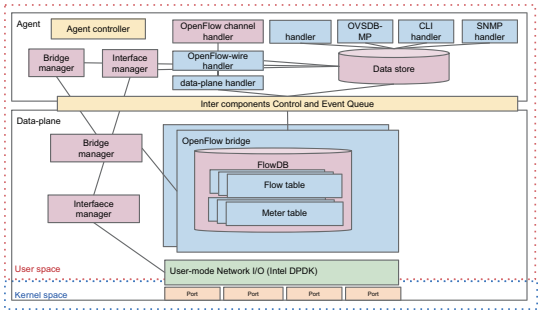

Figure 1: The design of the proposed software switch.

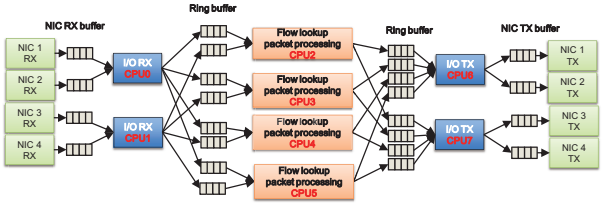

Figure 2: flow processing in data-plane

are performed with inter component control and event queuing framework. To prevent the overload of switch agent, the multi-criteria priority queueing are employed.

3 Implementation

The data-plane implementation leverages Intel DPDK library to accelerate network I/O performance by bypassing packet processing in OS kernel and direct access to NIC packet buffer from data-plane program. Figure 2 shows the implementation of flow processing in data-plane. To exploit processing power of many-core CPUs and to improve the efficiency of CPU cache and memory, the whole flow processing is decoupled to two type of processing: network I/O and packet processing (lookup and packet header modification). Current data-plane supports protocols that are often used in a carrier network, such as, MPLS, PBB, as well as protocols in a datacenter. The flow lookup in data-plane, which is most important part of software OpenFlow switch, employs the OpenFlow-aware modified Patricia-tree with flow lookup cache.

4 Evaluation

To conduct the basic performance of the switch, we employ a usecase of SDN/OpenFlow in wide-area network, a gateway function between a VLAN-based data center network and IP-VPN/MPLS. Figure 3 shows the performance evaluation setting. Figure 4 shows the one-way packet performance of the gateway with 100K flow rules in case of short packet (64Byte). The switch achieves 10.1MPPS (5.66Gbps) with 100K flow rules and short

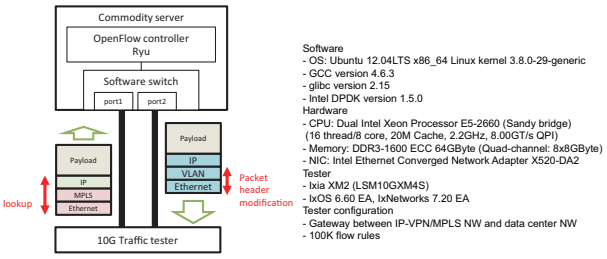

Figure 3: Performance evaluation environment of a gateway between IP-VPN/MPLS and data center

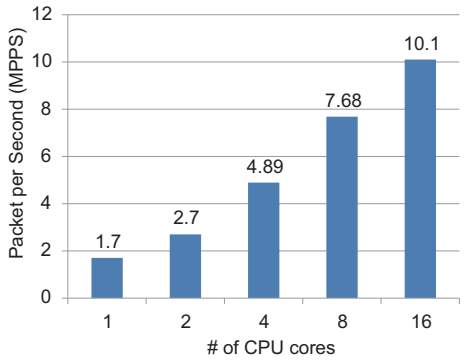

Figure 4: Packet processing performance of a gateway between IP-VPN/MPLS and data center with 100K flow rules.

packet (64B) using 16 CPUs. In case of long packet, the switch performed 10-Gbps wirerate.

Acknowledgments

This study is a part of the O_3 project supported by a grant of the Minitry of Internal Affairs and Communications in Japan.

References

[1] Fall, K. R., Iannaccone, G., Manesh, M., Ratnasamy, S., Argyraki, K. J., Dobrescu, M., and Egi, N. Routebricks: enabling general purpose network infrastructure. *Operating Systems Review 45*, 1 (2011), 112–125.

[2] McKeown, N., Anderson, T., Balakrishnan, H., Parulkar, G., Peterson, L., Rexford, J., Shenker, S., and Turner, J. Openflow: Enabling innovation in campus networks. *SIGCOMM Comput. Commun. Rev. 38*, 2 (Mar. 2008), 69–74.

[3] Pfaff, B., Pettit, J., Koponen, T., Amidon, K., Casado, M., and Shenker, S. Extending networking into the virtualization layer. In *Proc. of workshop on Hot Topics in Networks (HotNets-VIII)* (2009).

[4] SATO, Y., FUKUDA, I., and FUJITA, T. Deployment of openflow/sdn technologies to carrier services. *IEICE Transactions on Communications E96.B*, 12 (2013), 2946–2952.

[5] Zhou, D., Fan, B., Lim, H., Kaminsky, M., and Andersen, D. G. Scalable, high performance ethernet forwarding with cuckooswitch. In *CoNEXT* (2013), K. C. Almeroth, L. Mathy, K. Papagiannaki, and V. Misra, Eds., ACM, pp. 97–108.